

● 胡建平

谈机读目录的主题标引

ABSTRACT Subject search plays an important role in computer retrieval and also in MARC. Therefore, subject heading should be standardised according to *Classified Chinese Thesaurus* 16 refs

KEY WORDS Subject heading MARC

CLASS NUMBER G254 21

当前,我国各级各类图书馆纷纷运用电子计算机进行文献管理和文献检索,利用计算机进行主题检索已成为主要检索方式。而机读目录上的主题标引,则是开展主题检索的基础。

1 机读目录主题标引的重要性

主题词在机读目录中占有十分重要的位置,文献的主题标引工作又是保证机读目录质量的重要组成部分,也是开展主题检索的基础^[1]。

1.1 以主题标引为基础的主题检索已成为图书馆现代化的必然趋势

从文献的检索手段说,分类法和主题法都是揭示文献内容的检索语言,分类法是用人工标识的分类号来标识文献并按其学科知识体系组织起来,它能满足读者按科学体系检索图书的需要,但也存在着明显的不足:类目检索的复杂性、同主题图书的分散性、适用范围的局限性。主题检索语言是用规范化的自然语言来标识文献内容的主题,并将其按字顺组织起来,它可以随着学科的扩展和增加而不断地自我扩充,是一种开放性的体系。同分类法相比,它具有读者检索的直接性、反映主题的灵活性、同主题图书的集中性以及适用范围的广泛性^[2]。

文献检索的长期实践表明,主题检索所

占的比例大大高于其检索途径。曾有一份资料表明,对北图目录大厅做过目录使用的调查,发现主题目录的利用率是分类目录利用率的8倍,因为主题检索能适应学科交叉渗透的特点,不受学科的限制,读者不论需要哪一方面的文献,只要按相应的主题词去查找就能找到。这种检索文献的专指性和直观性,为不熟悉分类法及分类规则的读者提供了很大的方便。主题检索是一种读者检索方便的文献检索系统,是图书馆及信息部门的主要检索途径。

国际上许多大型的文献数据库都是通过主题词、关键词进行联机检索的,如果仍然只在现有的分类、著者、书名三套目录的基础上建立联机目录,很难和国际联机网络相适应,我们必须做好机读目录上的主题标引,建立主题检索系统,为国际国内的文献信息网络化打下基础。以主题标引为基础的主题检索是当今图书馆现代化发展的必然趋势。

1.2 以主题标引为基础的主题检索是图书馆标准化的需要

实现主题标引的标准化,要比实现分类标引标准化容易得多。分类标引是一种含有主观意志人为性的标识体系,由于各种原因,经常会出现同书不同号现象。我国两大中文图书统编系统,北京的CIP数据库统编系统和上海的SQLC数据库统编系统,由于其地

域性差异、专业性差异以及同类性差异的原因,在分析和选择类号的过程中会产生同书不同号的现象。例如:《行政领导与决策》一书,北京 CIP 系统将该书分入 D035 国家行政管理类;上海 SQLC 系统则分入 C933 领导学类。《当代中国人口发展与家庭户的变迁》一书,北京 CIP 系统将其分入 D669.1 社会生活与问题类;上海 SQLC 系统将其分入 C924.24 人口问题研究类^[3]。如采取主题标引的话,前书则会标引:领导决策-领导方法-决策学。后书则会标引:人口调查-中国;家庭问题-中国。主题标引能达到统一标引、统一检索的目的。无论各类型的图书馆在各时期所使用的分类法有多少差别,但一个国家只要有一部统一的主题词表,就可以达到统一标引、统一检索的目的。主题词表较少主观性和人为性。我们只采用一种通用语言,即主题标引的规范化,就可以达到统一检索的目的。以主题标引为基础的主题检索,是统一检索的需要,也是图书馆标准化的需要。

2 机读目录主题标引的标准化

书目数据标准化与规范化是保证书目数据的质量,形成书目数据网络化和实现资源共享的前提,其质量高低直接关系到数据库的存亡和兴衰,是数据库的生命线。一所图书馆的书目标准化程度,可以反映出其数据库的社会性和共用性。书目数据标准化程度高,就可以被其他图书馆或文献部门顺利地组织到自己的目录中去,有利于馆际、国际书目资源的交换,进而实现馆际、国际文献资源的共享。书目数据标准化是实现网络化的基础。

书目数据的标准化主要包括:文献著录、文献标引(分类标引和主题标引)、机读目录格式以及编目工作的规范控制等方面,本文则详细论述主题标引的标准化及它的规范控制问题。

2.1 主题标引必须以《中国分类主题词表》

为标准

所谓主题标引,是通过对文献的主题分析,并按照主题词表抽出代表该主题内容的主标题词的过程。主题标引必须以《中国分类主题词表》为标准。

《中国分类主题词表》是在《中国图书馆图书分类法》三版(含《中国图书资料法》三版)和《汉语主题词表》基础上编制的,它集分类、主题为一体,是一部大型的综合性分类法与叙词表对照索引式的一体化检索语言。它从内容到形式,从结构到语义等方面,实现了分类语言与主题语言、先组语言与后组语言的相互兼容^[4]。它的使用是我国分类主题一体化发展的又一重要标志^[5]。

《中国分类主题词表》有两大表系:“分类号—主题词对应表”和“主题词—分类号对应表”。

“分类号—主题词对应表”是以主题词注释《中国图书馆图书分类法》三版或《中国图书资料法》三版^[6],能从学科体系的角度全面展示本学科的主题词,在实现分类标引的同时,为主题标引选用较为专指的主题词提供方便。这样选词优点是专指性强,能较好地克服泛指词标引的错误。同时在相应类目下,对某些需要组配的学科、事物等,采用注释说明其组配标引方法,它的组配形式为各类型、各学科的文獻提供了不同的检索入口^[7]。

它的使用方法是対文献进行主题分析,确定该文献所论述的主题及其学科专业属性后,先标引分类号。依据分类号,找到该类目对应的主题词和主题词组配形式。在确定分类号的同时,获得主题标引。当类目所对应的主题表中找不到专指主题词时,则应采取归并列组配复分的方法,进一步选择专指词。

(1) 选用上位类目标引,对于该主题的内容的主题内容在主题词表中没有专指词,又不能组配表达,才可选用直接的上位主题词进行标引。例如:《珠江鱼类志》就可用上位类标引“水产志—珠江”。

(2) 从同位类目中补充, 如《异步化同步电机》, 分类标引为 TM 343, 对应的主题词是“异步电机”, 从同位类 TM 341 中选择“同步电机”; 组配标引为“异步电机—同步电机”。

(3) 依仿分规则, 从专类复分表、通用复分表中选词, 通过组配复分进一步缩小主题内容的范围。例如:《日本文化事业》, 分类标引为 G131.3, 对应主题词有文化事业, 从“世界地区表”中选择“日本”组配标引为“文化事业—日本”。

(4) 从相关类目中选词组配标引, 例如:《毛泽东与中国农业》分类标引为 A 841.695, 对应的主题词为“毛泽东思想”; 从相关类目选择“农业”; 组配标引为“毛泽东思想-农业-中国”。

“主题词—分类号对应表”相当于一部扩充了的《汉语主题词表》, 一部主题词型的《中国图书馆图书分类法》三版类目相关索引, 并将《中国图书馆图书分类法》的全部分类号置于相应的主题词或主题词串下^[8]。它能帮助标引人员就模棱两可的文献从主题标引角度实现分类标引, 对论述综合性主题的文献的分类标引起着指导作用。

它的使用方法是: 对文献进行主题分析, 确定主题概念, 将文献需要揭示和表达的主题概念转换为规范化的主题词或主题词串, 然后进行组配, 获得分类号。

“分类号—主题词对应表”是从分类角度标引文献的, “主题词—分类号对应表”是从主题角度标引文献的。一般来说, 对文献的标引, 以“分类号—主题词对应表”为主, 以“主题词—分类号对应表”为辅, 比较符合我国图书情报单位长期以来从分类的角度标引文献的情况^[9]。对论述综合性主题的文献, 适合用“主题词—分类号对应表”标引。

2.2 使用《中国分类主题词表》标引文献的意义

机读目录的主题标引, 以《中国分类主题词表》为标准, 能体现文献著录标准化的继承原则, 能利用计算机技术对主题标引进行规

范档控制, 以及能使分类检索系统和主题检索系统得到有机的统一, 促成分类检索与主题检索一体化发展。

(1) 使用《中国分类主题词表》标引图书文献, 体现了文献著录标准化的继承原则。

文献著录标准化, 应遵循的原则有互换原则、易于识别原则、转换原则与继承原则^[10]。我国各级、各类型图书馆, 原先大多采用《中国图书馆图书分类法》三版或《中国图书资料法》三版和《汉语主题词表》来类分、组织文献, 而《中国分类主题词表》是在它们的基础上编制的, 并将《中国图书馆图书分类法》的全部分类号置于相应的主题词或主题词串下, 它相当于一部扩充了的《汉语主题词表》。使用《中国分类主题词表》, 能使全国各级、各类型图书馆的文献标引, 在《中国分类主题词表》上统一起来, 并不打乱各馆原有的分类体系。使用《中国分类主题词表》标引文献, 体现了文献著录标准化的继承原则。

(2) 利用《中国分类主题词表》, 可以对主题标引进行规范档控制。

规范控制的问题, 是对同一检索点的标目出现前后不一致的问题而提出的。为了使数据处理达到文献检索点的前后一致性和惟一性, 在建立文献数据库的过程中, 必须对前后不一的现象做出统一规范, 对同一检索点的不同表示方式之间建立起联系。规范控制的核心要有规范档, 规范档是规范记录的有序集合^[11]。它记载了文献记录中选定标目的规范形式, 是著录人员按照条例选择的标目和确定的标目形式, 同时包括非规范标目, 用来说明规范标目的不同形式, 以及其他与规范标目有关的数据。笔者认为, 《中国分类主题词表》就是主题标目的规范档, 它可以对主题标引进行规范控制。规范控制是通过查询规范文档来维护文献记录标目的惟一过程。利用计算机技术, 把《中国分类主题词表》贮存在机内, 对机读目录的主题标引, 进行规范档控制。

(3) 使用《中国分类主题词表》标引图书

文献,能使分类检索系统和主题检索系统一体化。

主题检索有它的直接性、反映主题的灵活性以及便于特性检索等长处,但不便于族性检索;而分类检索也有它的系统性、符号的简便性以及便于族性检索的优点,但也存在不便于特性检索的缺点。电子计算机在图书馆及情报系统中的广泛应用,使建立一个大规模的具有系统分类和字顺主题两种途径的检索系统成为可能。在建立自动化检索系统的早期,情报界不少专家认为,分类法不能适应机器检索及多元检索的要求,但实践表明,只采用纯字顺的主题法的检索语言,是很难提高情报检索的检全率与检准率的^[12]。人们对检索语言表现概念之间关系的兴趣,已经从纯字顺的语词一览表,转向分类式的叙词表^[13]。

分类法与主题法是揭示主题内容、提供检索途径的不同方法,两者本来是一致的,最基本的数据元素(类目或主题词)在任何一个分类法或主题表中都是相同的,“人”只能是“人”,“学校”只能是“学校”。主题系统和分类系统都是运用同一原则^[14]。如果分类法附有主题索引,主题表附有范畴索引(即知识体系索引),互为补充,才能使分类主题一体化,才能使二者取长补短,互相融合,形成一个完整的系统。

《中国分类主题词表》是一体化的词表,它把分类法的系统性和主题法的直接性有机地结合在一起。分类主题一体化不仅仅限于分类表与主题词表相结合的一体化词表,而且还包括分类标引检索与主题标引检索的一体化^[15]。利用《中国分类主题词表》,可以一次标引同时获得分类标引和主题标引结果,这不仅提高了文献标引效率和质量,降低了标引难度,更重要的是能使两种各具特色的分类检索途径和主题检索途径、文献单元检索方式与标识单元检索方式以及分类检索系统与主题检索系统达到一体化。它能够把分类检索系统和主题检索系统有机地结合起

来,发挥两个检索系统的整体优势,它既能发挥分类检索工具的按学科组织文献,适应学科纵向发展,具有族性检索的优势,又能发挥主题检索工具的按事物的主题及其各个方面组织文献,适应跨学科或综合发展,具有特性检索的优势,形成一种新的检索语言^[16]。

参考文献

- 1 冯泽泗 机读目录的结构编制与应用 成都:成都科技大学出版社,1992:150
- 2 周子荣等 从分类目录到主题目录 见:文献编目论文选 北京:书目文献出版社,1992:58~66
- 3 徐菊兴,董勇敏 CIP 数据库与 SQLC 数据库系统差异性分析 图书情报工作,1997,(8)
- 4,7 贺定安 《中国分类主题词表》一体化标引方法研究 中国图书馆学报,1995,(6)
- 5,9 邓德生 论我国分类主题一体化研究的发展进程 大学图书馆学报,1996,(4)
- 6,8 刘湘生 《中国分类主题词表》使用说明 北京:华艺出版社,1994:6
- 10 李纪有 关于文献著录标准化的意义和原则 见:文献编目论文选 北京:书目文献出版社,1992:176~181
- 11 曾海燕 建立文献数据库规范控制的探讨 现代图书情报技术,1997,(1)
- 12 侯汉清 分类法的发展趋势简论 见:当代分类法主题法索引法研究 北京:书目文献出版社,1993:88~106
- 13 林德海译 美国图书馆学家和情报学家论现代图书分类法问题 见:中国图书馆图书分类法编辑会议资料 1979
- 14 陈源蒸 图书在版编目书目数据的标准化与规范化 北京:北京大学出版社,1994:93~156
- 15 郭锐 情报检索语言发展的重要方向——分类主题一体化 情报理论与实践,1989,(6)
- 16 肖自力等 加强分类目录,还是另编目录 见:文献编目论文选 北京:书目文献出版社,1992:67~78

胡建平 宁波大学邵逸夫图书馆工作。通讯地址:浙江省宁波市。邮编 315211。

(来稿时间:1997.11.25。编发者:赵薇)